

COMPUTATIONALLY EFFICIENT APPROXIMATIONS OF THE JOINT SPECTRAL RADIUS*

VINCENT D. BLONDEL[†] AND YURII NESTEROV[‡]

Abstract. The joint spectral radius of a set of matrices is a measure of the maximal asymptotic growth rate that can be obtained by forming long products of matrices taken from the set. This quantity appears in a number of application contexts but is notoriously difficult to compute and to approximate. We introduce in this paper a procedure for approximating the joint spectral radius of a finite set of matrices with arbitrary high accuracy. Our approximation procedure is polynomial in the size of the matrices once the number of matrices and the desired accuracy are fixed.

For the special case of matrices with nonnegative entries we give elementary proofs of simple inequalities that we then use to obtain approximations of arbitrary high accuracy. From these inequalities it follows that the spectral radius of matrices with nonnegative entries is given by the simple expression

$$\rho(A_1, \dots, A_m) = \lim_{k \rightarrow \infty} \rho^{1/k}(A_1^{\otimes k} + \dots + A_m^{\otimes k}),$$

where it is somewhat surprising to notice that the right-hand side does not directly involve any mixed product between the matrices. ($A^{\otimes k}$ denotes the k th Kronecker power of A .)

For matrices with arbitrary entries (not necessarily nonnegative), we introduce an approximation procedure based on semidefinite liftings that can be implemented in a recursive way. For two matrices, even the first step of the procedure gives an approximation whose relative accuracy is at least $1/\sqrt{2}$, that is, more than 70%. The subsequent steps improve the accuracy but also increase the dimension of the auxiliary problems from which the approximation can be found.

Our approximation procedures provide approximations of relative accuracy $1 - \epsilon$ in time polynomial in $n^{(\ln m)/\epsilon}$, where m is the number of matrices and n is their size. These bounds are close from optimality since we show that, unless $P=NP$, no approximation algorithm is possible that provides a relative accuracy of $1 - \epsilon$ and runs in time polynomial in n and $1/\epsilon$.

As a by-product of our results we prove that a widely used approximation of the joint spectral radius based on common quadratic Lyapunov functions (or on ellipsoid norms) has relative accuracy $1/\sqrt{m}$, where m is the number of matrices.

Key words. joint spectral radius, generalized spectral radius, approximation algorithm, switched systems, linear inclusion, Lyapunov indicator, Kronecker power

AMS subject classifications. 15A18, 15-04, 15A60, 90C22, 93-04

DOI. 10.1137/040607009

1. Introduction. Let $\|\cdot\|$ be a matrix norm. The *spectral radius* of the real matrix A is defined by

$$(1.1) \quad \rho(A) = \lim_{k \rightarrow +\infty} \|A^k\|^{1/k}.$$

*Received by the editors April 19, 2004; accepted for publication (in revised form) by L. Vandenberghe December 23, 2004; published electronically August 31, 2005. This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors. The research results presented in this paper have been supported in part by a grant "Action de recherche concertée ARC 04/09-315" from the "Direction de la recherche scientifique—Communauté française de Belgique."

<http://www.siam.org/journals/simax/27-1/60700.html>

[†]Division of Applied Mathematics, Université catholique de Louvain, 4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium (blondel@inma.ucl.ac.be).

[‡]Center for Operations Research and Econometrics (CORE), Université catholique de Louvain (UCL), 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium (nesterov@core.ucl.ac.be).

The spectral radius of a matrix does not depend on the chosen matrix norm, and it is also equal to $\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$ (see, for example, [22, Corollary 5.6.14]). The definition of spectral radius can be extended to *sets* of matrices in a natural way. The *joint spectral radius* of a set of matrices is a quantity introduced by Rota and Strang in the early 1960s that measures the maximal asymptotic growth rate that can be obtained by forming long products of matrices; see [29]. The general definition is for arbitrary sets of matrices, but we shall consider here only finite sets. Let $\{A_1, \dots, A_m\}$ be some set of real matrices. To the finite sequence $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_k) \in \{1, \dots, m\}^k$ we associate the corresponding matrix product

$$A_\sigma = A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}.$$

With this notation, the joint spectral radius is defined by

$$(1.2) \quad \rho(A_1, \dots, A_m) = \limsup_{k \rightarrow +\infty} \max_{\sigma \in \{1, \dots, m\}^k} \|A_\sigma\|^{1/k}.$$

As for the single-matrix case, the joint spectral radius does not depend on the matrix norm used. To see this, remember that any two matrix norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are related by $\alpha\|A\|_1 \leq \|A\|_2 \leq \beta\|A\|_1$ for some $0 < \alpha < \beta$. For any product $\sigma \in \{1, \dots, m\}^k$, one has $\alpha^{1/k}\|A_\sigma\|_1^{1/k} \leq \|A_\sigma\|_2^{1/k} \leq \beta^{1/k}\|A_\sigma\|_1^{1/k}$, and by letting k tend to infinity we conclude that the joint spectral radius is well defined independently of the matrix norm used.

A definition analogous to (1.2) is possible by replacing the norm appearing in the definition by a spectral radius. The quantity defined in this way is the *generalized spectral radius* introduced in [15]. In [5], the joint and generalized spectral radii of finite (or bounded) sets of matrices are proved to be equal (see also [35, Theorem 1] for an elementary proof), thus reinforcing the status of the joint spectral radius as a legitimate generalization of the spectral radius of a single matrix. In what follows, we shall deal only with the spectral radius defined with a norm, as in (1.2).

Since its introduction in the 1960s the joint spectral radius has appeared in a number of different contexts; see, e.g., [31] or [16] for recent short surveys.¹ Let us illustrate one application in a dynamical system context. Consider the simple discrete-time linear inclusion

$$x_{k+1} \in \{A_1, \dots, A_m\} x_k, \quad x_0 \in \mathbf{R}^n,$$

in which at each step a particular linear transformation is chosen from a finite number of possible choices. The maximal asymptotic rate of growth of the trajectories associated with such a system is given by the joint spectral radius $\rho(A_1, \dots, A_m)$. (In the discrete linear inclusion literature, the logarithm of the joint spectral radius is sometimes called *Lyapunov indicator*; see, for example, [2].) In particular, all possible trajectories will converge to the origin if and only if $\rho(A_1, \dots, A_m) < 1$. The joint spectral radius can thus be associated with the stability properties of time-varying linear systems in the worst case over all possible time variations. It also occurs in the context of “asynchronous” [33] or “desynchronized” [23] systems. Besides systems-theoretic interpretations, the concept is pervasive in many areas of applied mathematics such as in wavelets [15], iterated function systems, random walks, fractals, numerical solutions to ordinary differential equations [18], discrete-event systems [10], and coding theory [27].

¹Google returns 625 entries upon entry of the query “joint spectral radius.”

Despite its natural interpretation, the joint spectral radius is difficult to compute. Questions related to its computability and to the existence of efficient approximation algorithms have been posed more than a decade ago (see [34] and [24]). In principle, the spectral radius can be approximated to any desired accuracy by computing converging sequences of upper and lower bounds. The following bounds, proved in [24],

$$(1.3) \quad \max_{\sigma \in \{1, \dots, m\}^k} \rho(A_\sigma)^{1/k} \leq \rho(A_1, \dots, A_m) \leq \max_{\sigma \in \{1, \dots, m\}^k} \|A_\sigma\|^{1/k},$$

can be evaluated for increasing values of k and lead to arbitrary accurate approximations of ρ (see, e.g., [16] or [17]). Such approximation algorithms can in turn be used in procedures that decide, after finitely many steps, whether $\rho > 1$ or $\rho < 1$. Such procedures are given, e.g., by Brayton and Tong [13] in a system theory context and by Barabanov [2] in the context of discrete linear inclusions. These procedures may, however, not terminate when ρ happens to be equal to 1, and the existence of algorithms for computing arbitrarily precise approximations of ρ does therefore not rule out the possibility that the decision problem “ $\rho < 1$?” is undecidable. It is so far unknown whether this is the case or not (see [24] for a discussion of this issue and for a description of its connection with the finiteness conjecture that has since then been proved to be false; see [11] and [6]). A negative result in this direction is given in [8] where it is proved that the related problem “ $\rho \leq 1$?” is algorithmically undecidable.

An approach based on optimal control techniques for approximating the joint spectral radius is explained in [37] and in more detail in Appendix D of [14]. Approximations that are directly based on the inequalities (1.3) are expensive to compute. In [26], the exponential number of products that appear in the direct and naive computation of the bounds in (1.3) is reduced by avoiding duplicate computation of cyclic permutations; the total number of product to consider remains, however, exponential. In addition to this, there are to this date no known theoretical guarantees for the uniform rate of convergence of the bounds appearing in (1.3). Convergence rates are supplied in [36] and [37], but the bounds appearing there depend on the set of matrices considered. Approximations of arbitrary degree of accuracy can be computed but at a price that may happen to be prohibitive.

There are in fact intrinsic limitations for the rate at which the joint spectral radius can be approximated. Let us say that the value ξ approximates the value ρ with *relative accuracy* $\mu \in [0, 1]$ (or $100 \mu \%$) if $\mu \xi \leq \rho \leq \xi$. By using a small adaptation of a proof appearing in [32] we show that, unless $P=NP$, there is no algorithm that can compute the joint spectral radius of two matrices with relative accuracy $1 - \epsilon$ in time polynomial in the size of the matrices and in $1/\epsilon$ (see later for more precise definitions). Despite this negative result it is still conceivable, as pointed out in [9], that for any *fixed* desired accuracy, there exists a polynomial time algorithm that computes the joint spectral radius of the matrices with that accuracy. We prove in this paper that this is indeed the case. In Theorem 5, we prove that the joint spectral radius of m matrices of size n can be approximated with relative accuracy $1/\sqrt{m}$ by computing the spectral radius of a single matrix whose size is less than n^2 . This procedure can be applied in a recursive way, and in general we show how a relative accuracy of $(1/\sqrt{m})^{1/k}$ can be obtained by computing the spectral radius of a single matrix of size less than n^{2k} . As an illustration, the spectral radius of two matrices of size n can be computed with an accuracy of 70% by computing the spectral radius of a single matrix of size n^2 , and an accuracy of 95% can be obtained for the joint spectral radius of three matrices by computing the spectral radius of a single matrix of size n^{11} . More generally, for any number of matrices in the set and desired relative

accuracy we construct in section 4 a single matrix whose spectral radius approximates the joint spectral radius with the desired accuracy. The approximation procedure runs in polynomial time once the desired accuracy and the number of matrices in the set are fixed. More precisely, our approximation procedures provide approximations of relative accuracy $1 - \epsilon$ in time polynomial in $n^{(\ln m)/\epsilon}$, where m is the number of matrices and n is their size. Notice that $n^{(\ln m)/\epsilon} = e^{(\ln n \ln m)/\epsilon} = m^{(\ln n)/\epsilon}$, and so our approximation procedure also runs in time polynomial in the number of matrices once the desired accuracy and the size of the matrices are fixed. These bounds are close from optimality since we prove that, unless $P=NP$, no approximation polynomial time algorithm in n and $1/\epsilon$ is possible.

We now briefly describe how our results are obtained and how the paper is organized. We first notice in section 2 that for matrices with nonnegative entries the joint spectral radius satisfies

$$(1.4) \quad \frac{1}{m} \rho(A_1 + \cdots + A_m) \leq \rho(A_1, \dots, A_m) \leq \rho(A_1 + \cdots + A_m).$$

(The left-hand side inequality is valid for arbitrary matrices; there is no need to assume that the matrices have nonnegative entries.) Matrices with nonnegative entries are exactly those matrices A that are such that $A \mathbf{R}_+^n \subseteq \mathbf{R}_+^n$, and we prove in Theorem 1 that the inequalities (1.4) are satisfied not only for sets of matrices with nonnegative entries but also for sets of matrices that leave a proper cone invariant, i.e., matrices A_i that are such that $A_i K \subseteq K$ for some proper cone K and all i . Thus for these matrices the inequalities (1.4) provide a relative accuracy $1/m$. This accuracy can be improved by considering Kronecker powers of matrices. In section 3 we give an elementary proof that the joint spectral radius of the k th Kronecker powers of the matrices in a set is equal to the k th power of the joint spectral radius of the set (Theorem 3). Combining this with the relation (1.4) we prove that the approximation

$$\rho^{1/k}(A_1^{\otimes k} + \cdots + A_m^{\otimes k})$$

has relative accuracy $1/m^{1/k}$. Some of the arguments used to derive this result are valid only for matrices that have a common proper invariant cone. In section 4 we show how similar arguments can be used for arbitrary matrices. We introduce a semidefinite lifting procedure that transforms a linear operator acting \mathbf{R}^n into a linear operator acting on the space of symmetric $n \times n$ matrices. It appears that under this transformation the joint spectral radius is simply squared. Moreover, the lifting has the interesting feature that the operator defined in this way leaves the cone of semidefinite matrices invariant. This observation leads to a simple approximation procedure (Theorem 5) for arbitrary matrices. Implementation issues and numerical examples are provided in section 5. Our approximation procedure provides a relative accuracy $1 - \epsilon$ in time polynomial in $n^{(\ln m)/\epsilon}$. We prove in section 6 (Theorem 6) that the same accuracy cannot be obtained in time polynomial in n and $1/\epsilon$ unless $P=NP$. The semidefinite lifting introduced in section 4 is a fundamental and powerful tool for stability analysis. As an illustration of this, we prove in section 7 how to apply our results to analyze the quality of the so-called ellipsoid approximation of the joint spectral radius. We prove that the ellipsoid approximation (a notion that is formally defined in [6] but that is implicitly present in a number of earlier contributions on hybrid and time-varying systems; see, e.g., [21] and [1], as well as some of the references in [25]) is an approximation of guaranteed accuracy $1/\sqrt{m}$; a proof of this corollary can also be extracted from section 3 of [21]. It turns out that the ellipsoid approximation is also an approximation of a guaranteed accuracy $1/\sqrt{n}$ (see, e.g., [6] and [1]), and so

our bound is an improvement of this bound for the case where the dimension of the matrices is larger than the cardinality of the set. Finally, in a last section, we discuss some of our results.

2. Approximation for matrices leaving a cone invariant. In this section, we consider sets of matrices that leave a proper cone invariant, and we show with elementary arguments how joint spectral radius approximations of guaranteed accuracy can easily be computed for this case. We start with a proof that the spectral radius of a convex combination of matrices is always less than or equal to the joint spectral radius of the matrices. This result is valid for all sets of matrices (there is no need to assume that the matrices have nonnegative entries) and is proved in [7] using the main result of [23]. Here we present a direct and elementary justification.

LEMMA 1. *For any set of matrices $\{A_i : i = 1, \dots, m\}$ and any $\alpha_i \geq 0$ satisfying $\sum_{i=1}^n \alpha_i = 1$ we have*

$$(2.1) \quad \rho \left(\sum_{i=1}^m \alpha_i A_i \right) \leq \rho(A_1, \dots, A_m).$$

Proof. Let us fix some $\alpha_i \geq 0$ with $\sum_i \alpha_i = 1$ and an integer $k \geq 1$. Then

$$\left\| \left(\sum_{i=1}^m \alpha_i A_i \right)^k \right\| = \left\| \sum_{\sigma \in \{1, \dots, m\}^k} \alpha_\sigma A_\sigma \right\| \leq \sum_{\sigma \in \{1, \dots, m\}^k} \alpha_\sigma \|A_\sigma\| \leq \max_{\sigma \in \{1, \dots, m\}^k} \|A_\sigma\|,$$

but then also

$$\lim_{k \rightarrow \infty} \left\| \left(\sum_{i=1}^m \alpha_i A_i \right)^k \right\|^{1/k} = \limsup_{k \rightarrow \infty} \left\| \left(\sum_{i=1}^m \alpha_i A_i \right)^k \right\|^{1/k} \leq \limsup_{k \rightarrow \infty} \max_{\sigma \in \{1, \dots, m\}^k} \|A_\sigma\|^{1/k},$$

and the result then follows from the definitions of the spectral radius (1.1) and of the joint spectral radius (1.2). \square

An immediate consequence of this lemma is given by the following corollary.

COROLLARY 1.

$$\frac{1}{m} \rho \left(\sum_{i=1}^m A_i \right) \leq \rho(A_1, \dots, A_m). \quad \square$$

The example $A_1 = A, A_2 = -A$ clearly shows that we cannot hope in general to have $\rho(A_1, \dots, A_m) \leq \rho(\sum_{i=1}^m A_i)$. This inequality is nevertheless satisfied when the matrices A_i leave a proper cone invariant. A *cone* in \mathbf{R}^n is a subset $K \subseteq \mathbf{R}^n$ such that $\lambda v \in K$ for all $\lambda \geq 0$ and $v \in K$. We say that a cone K is *proper* if it is closed, convex, has nonempty interior, and contains no straight line. For example, the set of vectors with nonnegative entries \mathbf{R}_+^n is a proper cone, but \mathbf{R}^n itself is not (see [30] for more background on cones and proper cones). We shall say that the matrices A_i *leave a proper cone invariant* if there exists a proper cone $K \subseteq \mathbf{R}^n$ such that $A_i K \subseteq K$ for all i . For example, matrices with nonnegative entries leave the proper cone \mathbf{R}_+^n invariant.

LEMMA 2. *Associated with any proper cone K , there is a matrix norm $\|\cdot\|_K$ that satisfies $\|A\|_K \leq \|A+B\|_K$ for all matrices A and B that leave the cone K invariant.*

Proof. Most usual matrix norms satisfy this property when $K = \mathbf{R}_+^n$; we provide a construction for arbitrary proper cones K .

Let $\|\cdot\|$ be some arbitrary vector norm. The dual vector norm $\|\cdot\|_*$ is defined by $\|w\|_* = \max_{\|v\|=1} w^T v$. Assume that K is a proper cone. Define the dual of K by

$$K^* = \{w \in \mathbf{R}^n : w^T v \geq 0 \ \forall v \in K\}.$$

The dual of a proper cone is again a proper cone; see [30]. Let us now consider the quantity

$$\|A\|_K = \max_{\substack{v \in K, w \in K^* \\ \|v\| = \|w\|_* = 1}} w^T Av.$$

It is easy to verify that because K and its dual are proper cones, the quantity $\|\cdot\|_K$ is indeed a matrix norm. Moreover, by the definition of the norm, matrices A and B that satisfy $AK \subseteq K$ and $BK \subseteq K$ are such that $\|A\|_K \leq \|A + B\|_K$. \square

We may now state the result of this section.

THEOREM 1. *Let $\{A_i : i = 1, \dots, m\}$ be a set of matrices that leave a proper cone invariant. Then*

$$(2.2) \quad \frac{1}{m} \rho \left(\sum_{i=1}^m A_i \right) \leq \rho(A_1, \dots, A_m) \leq \rho \left(\sum_{i=1}^m A_i \right).$$

Proof. The lower bound in (2.2) is already established in Corollary 1. Assume that the matrices in $\{A_i : i = 1, \dots, m\}$ leave the proper cone K invariant, and define the matrix norm $\|\cdot\|_K$ as in Lemma 2. Since $A_i K \subseteq K$, we also have $A_\sigma K \subseteq K$ for all $\sigma \in \{1, \dots, m\}^k$, but then also

$$\max_{\sigma \in \{1, \dots, m\}^k} \|A_\sigma\|_K \leq \left\| \sum_{\sigma \in \{1, \dots, m\}^k} A_\sigma \right\|_K = \left\| \left(\sum_{i=1}^m A_i \right)^k \right\|_K,$$

and therefore

$$\begin{aligned} \rho(A_1, \dots, A_m) &= \limsup_{k \rightarrow \infty} \max_{\sigma \in \{1, \dots, m\}^k} \|A_\sigma\|_K^{1/k} \\ &\leq \limsup_{k \rightarrow \infty} \left\| \left(\sum_{i=1}^m A_i \right)^k \right\|_K^{1/k} = \rho \left(\sum_{i=1}^m A_i \right). \quad \square \end{aligned}$$

It is interesting to notice that the conclusion of Theorem 1 does not directly involve the cone K . The existence of a proper cone satisfying the hypotheses suffices to conclude, and the exact nature of the cone is irrelevant.

A matrix with nonnegative entries leaves the cone \mathbf{R}_+^n invariant, and so we have the following corollary.

COROLLARY 2. *Let the matrices A_i , $i = 1, \dots, m$, have nonnegative entries. Then*

$$\frac{1}{m} \rho \left(\sum_{i=1}^m A_i \right) \leq \rho(A_1, \dots, A_m) \leq \rho \left(\sum_{i=1}^m A_i \right). \quad \square$$

3. Kronecker lifting. We now describe a way to exploit Theorem 1 for obtaining approximations of arbitrary high accuracy. Our approximations involve Kronecker powers of matrices. The *Kronecker product* of two matrices $A \in R^{p_1 \times q_1}$ and $B \in R^{p_2 \times q_2}$ is the $p_1 p_2 \times q_1 q_2$ matrix defined by

$$A \otimes B = \begin{pmatrix} A_{1,1}B & \dots & A_{1,q_1}B \\ & \dots & \\ A_{p_1,1}B & \dots & A_{p_1,q_1}B \end{pmatrix}.$$

We also define the *Kronecker power* of a matrix:

$$A^{\otimes k} = \underbrace{A \otimes A \dots A \otimes A}_{k \text{ times}}.$$

There is no need for parentheses in this expression since the Kronecker product of two matrices is an associative operation. Let us prove some elementary properties.

LEMMA 3. *For matrices of appropriate sizes we have*

1. $(A \otimes B)^T = A^T \otimes B^T$;
2. $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$;
3. $(A^{\otimes k})^T = (A^T)^{\otimes k}$;
4. $A^{\otimes k} B^{\otimes k} = (AB)^{\otimes k}$;
5. *let $\|\cdot\|$ denote the spectral matrix norm induced by the standard Euclidean vector norm; then $\|A^{\otimes k}\| = \|A\|^k$.*

Proof. The first statement directly follows from the definition of the Kronecker product. The second statement is quite standard; see, e.g., section 4.2 in [22]. The third and fourth statement follow, respectively, from repeated application of the first and second statement. Finally, for proving the last statement, note that by using (3) and (4) we obtain

$$\|A^{\otimes k}\|^2 = \max_{\|x\|=1} x^T (A^{\otimes k})^T A^{\otimes k} x = \max_{\|x\|=1} x^T (A^T)^{\otimes k} A^{\otimes k} x = \max_{\|x\|=1} x^T (A^T A)^{\otimes k} x.$$

The matrix $(A^T A)^{\otimes k}$ is symmetric, and so the last expression is also equal to the largest magnitude of the eigenvalues of $(A^T A)^{\otimes k}$. For a matrix B the eigenvalues of the matrix $B^{\otimes k}$ are given by all possible products of k eigenvalues of B (see Theorem 4.2.12 in [22]), and we therefore conclude $\max_{\|x\|=1} x^T (A^T A)^{\otimes k} x = \|A\|^{2k}$. \square

We can now prove a useful identity for the spectral radius of Kronecker powers of matrices.

THEOREM 2. *Let $\{A_i : i = 1, \dots, m\}$ be a set of matrices and $l \geq 1$. Then*

$$\rho(A_1^{\otimes l}, \dots, A_m^{\otimes l}) = \rho^l(A_1, \dots, A_m).$$

Proof. Let $\sigma \in \{1, \dots, m\}^k$. By using the equalities (3) and (5) of Lemma 3 we obtain

$$\|A_{\sigma_k}^{\otimes l} \dots A_{\sigma_1}^{\otimes l}\| = \|A_{\sigma}^{\otimes l}\| = \|A_{\sigma}\|^l.$$

The result then follows from the definition of the joint spectral radius (1.2). \square

We now would like to combine Theorem 2 with Theorem 1 in order to obtain approximations of arbitrary high accuracy for matrices leaving a cone invariant. In order to do this we need to consider Kronecker products of cones. For two sets $Q_1 \subseteq R^n$ and $Q_2 \subseteq R^m$ let us denote

$$Q_1 \otimes Q_2 = \{z = u \otimes v : u \in Q_1, v \in Q_2\}.$$

(In this definition, the vectors u and v are treated as column matrices.) If K_1 and K_2 are proper cones, then $K_1 \otimes K_2$ does not need to be a proper cone. Indeed, consider, for example, the proper cones $K_1 = K_2 = \mathbf{R}_+^2$ for which it is easy to see that $K_1 \otimes K_2$ has empty interior in \mathbf{R}^4 . We do, however, have the following result.

LEMMA 4. *Let $K_1 \subset \mathbf{R}^n$ and $K_2 \subset \mathbf{R}^m$ be proper cones; then the cone $K = \overline{\text{Conv}(K_1 \otimes K_2)} \subset \mathbf{R}^{nm}$ is also a proper cone.*

Proof. The cone K is closed and convex by definition. Assume that $\text{int } K = \emptyset$. Then there exists $A \in \mathbf{R}^{nm} \setminus \{0\}$ such that

$$\langle A, z \rangle = 0 \quad \forall z \in K \supset K_1 \otimes K_2.$$

Note that the function $f(u, v) = \langle A, u \otimes v \rangle$ is a bilinear form defined on $K_1 \times K_2$. Since $\text{int}(K_1 \times K_2) \neq \emptyset$, f can vanish on this set only if $A = 0$. That is a contradiction.

Let us prove now that K contains no straight line. Assume that this is not the case. Then $\text{int } K^* = \emptyset$, which implies the existence of some $B \in \mathbf{R}^{nm} \setminus \{0\}$ such that

$$(3.1) \quad \langle B, w \rangle = 0 \quad \forall w \in K^*.$$

Since $\langle u \otimes v, x \otimes y \rangle = \langle u, x \rangle \cdot \langle v, y \rangle$, we have $K_1^* \otimes K_2^* \subseteq K^*$. Since K_1^* and K_2^* have nonempty interior, relation (3.1) is impossible by the first part of the proof. \square

We are ready to state the main result of this section.

THEOREM 3. *Let $\{A_i : i = 1, \dots, m\}$ be a set of matrices leaving a proper cone invariant. Then*

$$\frac{1}{m^{1/k}} \rho^{1/k}(A_1^{\otimes k} + \dots + A_m^{\otimes k}) \leq \rho(A_1, \dots, A_m) \leq \rho^{1/k}(A_1^{\otimes k} + \dots + A_m^{\otimes k}).$$

In particular, the joint spectral radius is given by

$$(3.2) \quad \rho(A_1, \dots, A_m) = \lim_{k \rightarrow \infty} \rho^{1/k}(A_1^{\otimes k} + \dots + A_m^{\otimes k}).$$

Proof. Assume that the matrices leave the proper cone K invariant. If $AK \subseteq K$, then $(AK)^{\otimes k} \subseteq K^{\otimes k}$ and, by Lemma 3, $A^{\otimes k}K^{\otimes k} \subseteq K^{\otimes k}$. But then also $A^{\otimes k} \overline{\text{Conv } K^{\otimes k}} \subseteq \overline{\text{Conv } K^{\otimes k}}$. By Lemma 4 the cone $\overline{\text{Conv } K^{\otimes k}}$ is a proper cone. Thus the matrices $A_i^{\otimes k}$ leave a proper cone invariant, and by Theorem 1 we have

$$\frac{1}{m} \rho(A_1^{\otimes k} + \dots + A_m^{\otimes k}) \leq \rho(A_1^{\otimes k}, \dots, A_m^{\otimes k}) \leq \rho(A_1^{\otimes k} + \dots + A_m^{\otimes k}).$$

In order to conclude it suffices to use the fact, proved in Theorem 2, that

$$\rho(A_1^{\otimes k}, \dots, A_m^{\otimes k}) = \rho^k(A_1, \dots, A_m). \quad \square$$

It is somewhat surprising to notice that the right-hand side in (3.2) does not directly involve any mixed product between the matrices.

An approximation of relative accuracy $(1/m)^{1/k}$ can thus be obtained by computing the spectral radius of a single matrix of dimension n^k . For pairs of matrices, some of the relative accuracies and the corresponding matrix sizes are as follows:

Relative accuracy	0.707	0.840	0.917	0.957
Matrix size	n^2	n^4	n^8	n^{16}

4. Semidefinite lifting. The result in Theorem 3 provides an easy way to evaluate with arbitrary accuracy the joint spectral radius of matrices that leave a proper cone invariant. Unfortunately, matrices do not always leave a proper cone invariant. We show in this section how an invariant cone can always be created by semidefinite lifting.

Let $A \in \mathbf{R}^{n \times n}$ and consider the following linear operator:

$$(4.1) \quad X \rightarrow AXA^T : \mathbf{R}^{n \times n} \rightarrow \mathbf{R}^{n \times n}.$$

A matrix representation for this linear operator can be obtained by using the matrix-to-vector operator that develops a matrix into a vector by taking its columns one by one. This operator, denoted vec , satisfies the elementary property $\text{vec}(CXD) = (D^T \otimes C) \text{vec}(X)$ (see Lemma 4.3.1 in [22]). We therefore have $\text{vec}(AXA^T) = (A \otimes A) \text{vec}X$, and a matrix representation of the linear operator (4.1) is thus simply given by $A \otimes A$.

Consider now the space of *symmetric* matrices \mathbf{S} . This space is a subspace of $\mathbf{R}^{n \times n}$ of dimension $n(n + 1)/2$, and the operator

$$(4.2) \quad X \rightarrow AXA^T : \mathbf{S} \rightarrow \mathbf{S}$$

is again a linear operator. Matrix representations of this operator are of course different from those of the operator (4.1) since, in particular, the spaces on which these operators operate have different dimensions. (We describe in the next section how to construct a matrix representation for the operator on symmetric matrices.) In the next theorem we prove that even though the operators $X \rightarrow AXA^T$ on \mathbf{S} and on $\mathbf{R}^{n \times n}$ are different, their joint spectral radius are equal.

THEOREM 4. *Let $\{A_i \in \mathbf{R}^{n \times n} : i = 1, \dots, m\}$, and denote by M_A a matrix representation of the linear operator $X \rightarrow AXA^T : \mathbf{S} \rightarrow \mathbf{S}$. Then we have*

$$\rho(M_{A_1, \dots, A_m}) = \rho(A_1 \otimes A_1, \dots, A_m \otimes A_m) = \rho^2(A_1, \dots, A_m).$$

Proof. The second equality is already proved in Theorem 2; we need only to prove the first equality. Let $\|\cdot\|_F$ denote the Frobenius matrix norm (i.e., the square root of the sum of squares of all entries), and consider the resulting induced operator norms for $X \rightarrow AXA^T$ on $\mathbf{R}^{n \times n}$ and \mathbf{S} :

$$\sup_{X \in \mathbf{R}^{n \times n}, \|X\|_F=1} \|AXA^T\|_F \quad \text{and} \quad \sup_{X \in \mathbf{S}, \|X\|_F=1} \|AXA^T\|_F.$$

We claim that these two operator norms are equal and that the supremum is achieved for some symmetric matrix of rank one. Indeed, using the matrix-to-vector operator vec and denoting by $\|\cdot\|$ the usual vector Euclidean norm we get

$$\begin{aligned} \sup_{X \in \mathbf{R}^{n \times n}, \|X\|_F=1} \|AXA^T\|_F^2 &= \sup_{x \in \mathbf{R}^{n^2}, \|x\|=1} \|(A \otimes A)x\|^2 \\ &= \sup_{x \in \mathbf{R}^{n^2}, x^T x=1} x^T (A \otimes A)^T (A \otimes A)x \\ &= \sup_{x \in \mathbf{R}^{n^2}, x^T x=1} x^T (A^T A \otimes A^T A)x. \end{aligned}$$

For deriving the last equality, we have used point 2 in Lemma 3. The supremum in the last expression is achieved by any eigenvector x associated with an eigenvalue

of largest magnitude of $A^T A \otimes A^T A$. Let $(\lambda_i, v_i), i = 1, \dots, n$, be the set of eigenvalues/eigenvectors pairs associated with the symmetric matrix $A^T A$, and let λ_1 be such that $|\lambda_1| \geq \lambda_i$ for all i . The eigenvalues/eigenvectors pairs of the matrix $A^T A \otimes A^T A$ are then given by $(\lambda_i \lambda_j, v_i \otimes v_j)$ for $i, j = 1, \dots, n$. The vector $v_1 \otimes v_1$ is thus an eigenvector of $A^T A \otimes A^T A$ of largest eigenvalue magnitude. Since $\text{vec}(v_k v_k^T) = v_k \otimes v_k$ we see that the supremum of $\|AXA^T\|_F$ is achieved for the symmetric rank one matrix $v_1 v_1^T$. From this we conclude that

$$\sup_{X \in \mathbf{R}^{n \times n}, \|X\|_F=1} \|AXA^T\|_F = \sup_{X \in \mathbf{S}, \|X\|_F=1} \|AXA^T\|_F,$$

and thus also

$$\sup_{\|x\|=1} \|(A \otimes A)x\| = \sup_{\|x\|=1} \|(M_A)x\|.$$

Notice that $M_{A_\sigma} = (M_{A_i})_\sigma$, and so

$$\|(M_{A_i})_\sigma\| = \|M_{A_\sigma}\| = \|A_\sigma \otimes A_\sigma\| = \|(A_i \otimes A_i)_\sigma\|;$$

it then suffices to apply the definition of the joint spectral radius to conclude the proof. \square

A matrix A is *positive semidefinite* (which we denote by $A \succeq 0$) if $v^T A v \geq 0$ for all $v \in \mathbf{R}^n$. The set of symmetric positive semidefinite matrices, denoted \mathbf{S}_+ , is a cone. It is not a proper cone of $\mathbf{R}^{n \times n}$ because it has empty interior in $\mathbf{R}^{n \times n}$, but it is a proper cone of the set of symmetric matrices \mathbf{S} . If $X \in \mathbf{S}_+$, then we clearly have $AXA^T \in \mathbf{S}_+$, and so the operator $X \rightarrow AXA^T$ leaves the proper cone \mathbf{S}_+ invariant. Combining this observation with Theorems 3 and 4 we deduce the following theorem.

THEOREM 5. *Let $\{A_i \in \mathbf{R}^{n \times n} : i = 1, \dots, m\}$, and denote by M_A a matrix representation of the linear operator $X \rightarrow AXA^T : \mathbf{S} \rightarrow \mathbf{S}$. Then*

$$(4.3) \quad \frac{1}{\sqrt{m}} \rho^{1/2}(M_{A_1} + \dots + M_{A_m}) \leq \rho(A_1, \dots, A_m) \leq \rho^{1/2}(M_{A_1} + \dots + M_{A_m}).$$

Note that for $m = 2$ the bounds (4.3) are not so bad. Indeed, $1/\sqrt{2} > 0.7$, and so the relative accuracy of the approximation is at least 70%. The main interest of this result resides, however, in the possibility of applying it recursively. Starting from an initial matrix $A \in \mathbf{R}^{n \times n}$, we define a sequence of operators acting on spaces of increasing dimensions. For $l = 1$ we define

$$A^{[1]}(x) \equiv Ax, \quad x \in E_1 \equiv \mathbf{R}^n,$$

and for $l \geq 2$ we define the operators recursively:

$$A^{[l]}(x) \equiv M_{A^{[l-1]}}(x), \quad x \in E_l \equiv \mathbf{S}(E_{l-1}).$$

Assume for the simplicity of the presentation that $m = 2$, and let $A_1, A_2 \in \mathbf{R}^{n \times n}$. From Theorem 4 we know that

$$\rho(A_1^{[l]}, A_2^{[l]}) = \rho^{2^l}(A_1, A_2).$$

On the other hand, for any $l \geq 1$ the operators $A_1^{[l]}$ and $A_2^{[l]}$ leave the cone $\mathbf{S}_+(E_{l-1})$ invariant. Therefore, in view of Theorem 1, we have

$$\frac{1}{2} \rho(A_1^{[l]} + A_2^{[l]}) \leq \rho(A_1^{[l]}, A_2^{[l]}) \leq \rho(A_1^{[l]} + A_2^{[l]}).$$

Combining these last two expressions, we get the following bounds:

$$\left(\frac{1}{2}\right)^{1/2^l} \left[\rho\left(A_1^{[l]} + A_2^{[l]}\right)\right]^{1/2^l} \leq \rho(A_1, A_2) \leq \left[\rho\left(A_1^{[l]} + A_2^{[l]}\right)\right]^{1/2^l}.$$

Note that

$$1 - \frac{\ln 2}{2^l} \leq \left(\frac{1}{2}\right)^{1/2^l} \leq 1,$$

and thus the improvement in the quality of our approximation is quite fast. Unfortunately, the dimensions of the spaces E_l are also growing fast; we have

$$n_{l+1} = \frac{1}{2}n_l(n_l + 1)$$

(where $n_i = \dim E_i$), and therefore asymptotically we have $n_l = O((n/2)^{2^l})$. Let us display, for a pair of matrices, the relative accuracy of our approximation as a function of the resulting dimension.

Steps	Accuracy	$n = 2$	$n = 10$	$n = 100$
1	0.707	3	55	5050
2	0.840	6	1540	*
3	0.917	21	118570	*
4	0.957	231	*	*
5	0.978	26796	*	*

We use the symbol * to mark the cases for which the dimension of the auxiliary problem goes beyond the abilities of modern computers.

5. Numerical implementation and examples. The recursive semidefinite lifting for obtaining approximations of the joint spectral radius of increasing accuracy may be difficult to implement because the matrix M_A is not easy to express in terms of the matrix A . Consider, for example, the case of 2×2 matrices. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix};$$

then one can easily compute

$$M_A = \begin{pmatrix} a_{11}^2 & 2a_{11}a_{12} & a_{12}^2 \\ a_{11}a_{21} & a_{11}a_{22} + a_{12}a_{21} & a_{12}a_{22} \\ a_{21}^2 & 2a_{21}a_{22} & a_{22}^2 \end{pmatrix}.$$

In general, the matrix M_A can be expressed as follows:

$$(5.1) \quad M_A = (Q^T Q)^{-1} Q^T P^{-1} (A \otimes A) P Q,$$

where P is a particular permutation matrix of size n^2 and Q is given by

$$Q = \begin{pmatrix} I_n & 0 \\ 0 & I_{\frac{n(n-1)}{2}} \\ 0 & I_{\frac{n(n-1)}{2}} \end{pmatrix}.$$

(The matrix I_k is the identity matrix of size k .) One difficulty with expression (5.1) is that the permutation matrix P is tedious to construct; it essentially relates the column decomposition of a matrix with its diagonal decomposition. The construction

of the permutation matrix is somewhat cumbersome, but it can be automated. A way to bypass repeated computation of the matrix for larger and larger dimensions can be achieved by using the semidefinite lifting only once and then applying only Kronecker liftings. The semidefinite lifting can be performed either at the beginning or at the end of the process, leading to the bounds

$$(5.2) \quad \frac{1}{m} \rho(M_{A_1}^{\otimes l} + \dots + M_{A_m}^{\otimes l}) \leq \rho^{2l}(A_1, \dots, A_m) \leq \rho(M_{A_1}^{\otimes l} + \dots + M_{A_m}^{\otimes l})$$

and

$$(5.3) \quad \frac{1}{m} \rho(M_{A_1^{\otimes l}} + \dots + M_{A_m^{\otimes l}}) \leq \rho^{2l}(A_1, \dots, A_m) \leq \rho(M_{A_1^{\otimes l}} + \dots + M_{A_m^{\otimes l}}).$$

The validity of these inequalities results from the combination of Theorem 1, Theorem 2, Lemma 4, and Theorem 4. Expressions (5.2) and (5.3) both provide a relative accuracy of $(1/m)^{1/(2l)}$ but involve matrices of different size. The matrices $M_{A_i}^{\otimes l}$ have size $(n(n+1)/2)^l$, whereas the matrices $M_{A_i^{\otimes l}}$ have size $n^l(n^l+1)/2 > (n(n+1)/2)^l$. In addition to this, the matrices $M_{A_i}^{\otimes l}$ are easier to compute than $M_{A_i^{\otimes l}}$ since they necessitate the evaluation of semidefinite liftings of smaller matrices. The approximation based on (5.2) is therefore to be preferred because it is easier to evaluate and leads to matrices of smaller dimension.

Of course, other combinations of Kronecker and semidefinite liftings are also possible. Whenever a k th Kronecker power is used, the current relative accuracy is k th rooted, and whenever a semidefinite lifting is used, the relative accuracy is square rooted. Moreover, at least one semidefinite lifting is needed if the original matrices do not leave a proper cone invariant.

For the numerical implementation of these approximations, notice also that the spectral radius of $M_{A_1} + \dots + M_{A_m}$ is the spectral radius of the linear operator

$$B : X \rightarrow A_1 X A_1^T + \dots + A_m X A_m^T : \mathbf{S} \rightarrow \mathbf{S},$$

and it can therefore be found by standard linear algebra techniques or by solving the following linear inequality optimization problem,

$$(5.4) \quad \inf_{X, \tau} \{ \tau : \tau X \succeq A_1 X A_1^T + \dots + A_m X A_m^T, X \succ 0 \},$$

for which efficient convex optimization algorithms are available (see, e.g., [12]) that are implemented in standard commercial softwares such as the MATLAB “LMI Control Toolbox.”

6. Computational complexity analysis. The joint spectral radius can be approximated to arbitrary accuracy. It is proved in [32] that, unless P=NP, approximating algorithms of relative accuracy $1 - \epsilon$ for pairs of matrices cannot run in time polynomial in the size of the matrices and in $\ln(1/\epsilon)$. It was later noticed (see note 9 on page 1260 of [9]) that a careful examination of the proof in [32] shows that, unless P=NP, there are no approximation algorithms of relative accuracy $1 - \epsilon$ that run in time polynomial in the size of the matrices and in $1/\epsilon$. This is, however, not proved in [9]. In this section we prove this result by extracting the essential part of the proof in [32].

We proceed by reduction from the classical NP-complete satisfiability problem 3-SAT. This problem is defined as follows. Consider a set $\{x_1, \dots, x_n\}$ of Boolean variables. A *literal* is either a Boolean variable, x_i , or its negation, \bar{x}_i . A *three-literal*

clause is a disjunction of three literals (e.g., $x_3 \vee \bar{x}_5 \vee \bar{x}_6$). The 3-SAT problem is the problem of determining, for a given collection of three-literal clauses, if there exists a truth assignment for the variables that simultaneously satisfies all clauses. This problem is known to be NP-complete; finding a polynomial time algorithm for the problem is equivalent to proving that $P=NP$.

The complexity result proved in [32] is based on the construction of a pair of matrices which we briefly outline. (The complete construction can easily be extracted from [32].) From a given instance of 3-SAT with n variables and p clauses, two square matrices A_1, A_2 are constructed in polynomial time. The matrices have their entries in $\{0, 1\}$, are of size $(n + 1)(p + 1)$, and have a joint spectral radius that satisfies

$$\begin{aligned} \rho(A_1, A_2) &\geq p^{1/(n+2)} && \text{if the instance is satisfiable,} \\ \rho(A_1, A_2) &\leq (p - 1)^{1/(n+2)} && \text{if the instance is not satisfiable.} \end{aligned}$$

Any algorithm of relative accuracy $(1 - 1/p)^{1/(n+2)}$ allows us to make the distinction between these two cases. Moreover, since

$$\left(1 - \frac{1}{p}\right)^{1/(n+2)} \leq 1 - \frac{1}{p(n+2)} \leq 1,$$

it is clear that an approximation algorithm of relative accuracy $1 - 1/p(n+2)$ also allows us to decide 3-SAT. Since the size of the matrices in the construction are given by $(n + 1)(p + 1)$, we deduce the following theorem.

THEOREM 6. *Unless $P=NP$, the problem of approximating the joint spectral radius of two square matrices with $\{0, 1\}$ entries and with relative accuracy $1 - \epsilon$ cannot be obtained in time polynomial in the size of the matrices and in $1/\epsilon$.*

From Theorem 3 we know that the spectral radius of two matrices with nonnegative entries is given by $\rho(A_1, A_2) = \lim_{k \rightarrow \infty} \rho^{1/k}(A_1^{\otimes k} + A_2^{\otimes k})$, and so the quantity $\lim_{k \rightarrow \infty} \rho^{1/k}(A_1^{\otimes k} + A_2^{\otimes k})$ is NP-hard to approximate in the sense given in Theorem 6. Theorem 3 provides also a rate of convergence for the approximations $\rho^{1/k}(A_1^{\otimes k} + A_2^{\otimes k})$. From this we may prove the following complexity result.

THEOREM 7. *The problem of determining, for a given integer $k \geq 0$ and for a given pair of matrices A_1, A_2 with nonnegative entries, if*

$$\rho(A_1^{\otimes k} + A_2^{\otimes k}) < 1$$

is a problem that is NP-hard.

Proof. We proceed by reduction from 3-SAT. From a given instance of 3-SAT with n variables and p clauses we use the construction given in [32] to construct two square matrices B_1, B_2 that have their entries in $\{0, 1\}$, are of size $(n + 1)(p + 1)$, and have a joint spectral radius that satisfies

$$\begin{aligned} \rho(B_1, B_2) &\geq p^{1/(n+2)} && \text{if the instance is satisfiable,} \\ \rho(B_1, B_2) &\leq (p - 1)^{1/(n+2)} && \text{if the instance is not satisfiable.} \end{aligned}$$

We then choose $\alpha \in \mathbf{R}$ with $0 < \alpha < 1$, $r \in \mathbf{Q}$, and $k \geq 0$ such that $\alpha < (1/2)^{1/k}$ and

$$(6.1) \quad (p - 1)^{\frac{1}{n+2}} < \frac{1}{\alpha}(p - 1)^{\frac{1}{n+2}} < r < \alpha q^{\frac{1}{n+2}} < p^{\frac{1}{n+2}}.$$

Consider now the matrices $A_1 = (1/r)B_1$ and $A_2 = (1/r)B_2$. We claim that $\rho(A_1^{\otimes k} + A_2^{\otimes k}) < 1$ if and only if the instance of 3-SAT is not satisfiable. Indeed, assume first that the instance of 3-SAT is not satisfiable. Then

$$\rho(B_1, B_2) \leq (p - 1)^{1/(n+2)}.$$

Since $\rho(A_1^{\otimes k} + A_2^{\otimes k})$ is an approximation of relative accuracy $1/2^{2k}$ of $\rho(A_1, A_2)$, we have

$$\rho(A_1^{\otimes k} + A_2^{\otimes k}) \leq 2^{2k} \rho(A_1, A_2) \leq \frac{2^{2k}}{r} \rho(B_1, B_2) \leq \frac{2^{2k}}{r} (p-1)^{1/(n+2)},$$

and by using the inequalities (6.1) we conclude that $\rho(A_1^{\otimes k} + A_2^{\otimes k}) < 1$. In an analogous way one can see that if the instance of 3-SAT is satisfiable, then $\rho(A_1^{\otimes k} + A_2^{\otimes k}) > 1$, and the proof is therefore complete. \square

7. Ellipsoid approximation. Let us now compare our approximation of the joint spectral radius with another approximation appearing in the literature. The following approximation, called ellipsoid approximation, is introduced in [1] (see also [7]):

$$(7.1) \quad \hat{\rho}(A_1, \dots, A_m) = \left[\inf_{X, \tau} \{ \tau X \succeq A_i X A_i^T, i = 1, \dots, m, X \succ 0 \} \right]^{1/2}.$$

This expression corresponds to the best ellipsoid norm for the set of matrices. In [7] the following inequalities are proved:

$$\frac{1}{\sqrt{n}} \hat{\rho}(A_1, \dots, A_m) \leq \rho(A_1, \dots, A_m) \leq \hat{\rho}(A_1, \dots, A_m).$$

Thus, the ellipsoid approximation has a relative accuracy of $1/\sqrt{n}$. This accuracy decreases with the size of the matrices but does not depend on the number of matrices in the set. Using our results we can prove that the relative accuracy of the ellipsoid approximation is in fact bounded by $1/\sqrt{m}$, where m is the number of matrices.

THEOREM 8. *Let $\{A_i : i = 1, \dots, m\}$ be a set of matrices, and define the ellipsoid approximation $\hat{\rho}$ by (7.1). Then*

$$\frac{1}{\sqrt{m}} \hat{\rho}(A_1, \dots, A_m) \leq \rho(A_1, \dots, A_m) \leq \hat{\rho}(A_1, \dots, A_m).$$

Proof. It is easy to see that $\rho(A_1, \dots, A_m) \leq \hat{\rho}(A_1, \dots, A_m)$. For proving the first inequality, note that the spectral radius of the linear operator

$$B : X \rightarrow \sum_{i=1}^m A_i X A_i^T : \mathbf{S} \rightarrow \mathbf{S}$$

can be represented as follows:

$$(7.2) \quad \rho(B) = \inf_{X, \tau} \left\{ \tau : \tau X \succeq \sum_{i=1}^m A_i X A_i^T, X \succ 0 \right\}.$$

If a pair (X, τ) is feasible for the optimization problem in (7.2), then it is also feasible for the optimization problem in (7.1). Therefore

$$\hat{\rho}(A_1, \dots, A_m) \leq \rho^{1/2}(B).$$

Finally, by Theorem 5 we have

$$\frac{1}{\sqrt{m}} \rho^{1/2}(B) \leq \rho(A_1, \dots, A_m),$$

and so we get the following bounds:

$$(7.3) \quad \frac{1}{\sqrt{m}}\hat{\rho}(A_1, \dots, A_m) \leq \frac{1}{\sqrt{m}}\rho^{1/2}(B) \leq \rho(A_1, \dots, A_m) \leq \hat{\rho}(A_1, \dots, A_m) \leq \rho^{1/2}(B).$$

Thus the ellipsoid estimate $\hat{\rho}(A_1, \dots, A_m)$ has relative quality $1/\sqrt{m}$. \square

By using Theorem 2 one can obtain a statement analogous to that of Theorem 3.

COROLLARY 3. *Let $\{A_i : i = 1, \dots, m\}$ be a set of matrices, and define the ellipsoid approximation $\hat{\rho}$ by (7.1). Then*

$$(7.4) \quad \frac{1}{m^{1/(2k)}} \hat{\rho}^{1/k}(A_1^k, \dots, A_m^k) \leq \rho(A_1, \dots, A_m) \leq \hat{\rho}^{1/k}(A_1^k, \dots, A_m^k).$$

For $k = 1$, this result is also proved in [1]. For arbitrary k , a proof can be reconstructed from the result proved in [21]. Notice, however, that approximation algorithms that are directly based on (7.4) are not efficient because they require solving linear matrix inequalities of very large dimension. Approximations based on (5.2) are easier to obtain because there exist many efficient numerical algorithms for approximating the spectral radius of a matrix.

8. Discussion. The results presented in this paper introduce the possibility of computing approximations for the joint spectral radius with worst-case theoretical guarantees. Of course, the computational complexity of these estimates needs further examination. Indeed, the spectral structure of the matrix $A^{\otimes k}$ is very simple, and this simplicity must be inherited somehow by the matrix $A_1^{\otimes k} + A_2^{\otimes k}$. It therefore appears to be an interesting problem to investigate the possibility of constructing efficient procedures for finding the spectral radius of this sum; of course there are theoretical limitations on what can be achieved since we have shown that the problem of deciding $\rho(A_1^{\otimes k} + A_2^{\otimes k}) \leq 1$ is NP-hard.

Another interesting issue is that of determining if worst-case theoretical guarantees can be provided for the approximation

$$\max_{\sigma \in \{1, \dots, m\}^k} \rho(A_\sigma)^{1/k} \leq \rho(A_1, \dots, A_m).$$

In many numerical examples, this approximation does in fact perform at least as well.

Remark. When the paper was ready for publication, the authors discovered a very interesting theory developed by Protasov [28]. This theory also admits an interpretation in terms of Kronecker powers of matrices. However, it is focused mainly on geometrical properties and objects (invariant sets, p -sums of convex sets, etc.). Our approach is different in several aspects. First, it is simpler and it directly focused on computational aspects. Second, for some important cases (matrices with nonnegative elements) we have finer results. Finally, for our purposes we do not need additional assumptions (irreducibility).

REFERENCES

- [1] T. ANDO AND M.-H. SHIH, *Simultaneous contractibility*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 487–498.
- [2] N. E. BARABANOV, *On the Lyapunov exponent of discrete inclusions*. I, Avtomat. i Telemekh., no. 2, 1988, pp. 40–46; translation in *Automat. Remote Control*, 49 (1988), pp. 152–157.
- [3] N. E. BARABANOV, *On the Lyapunov exponent of discrete inclusions*. II, Avtomat. i Telemekh., no. 3, 1988, pp. 24–29; translation in *Automat. Remote Control*, 49 (1988), pp. 283–287.
- [4] N. E. BARABANOV, *On the Lyapunov exponent of discrete inclusions*. III, Avtomat. i Telemekh., no. 5, 1988, pp. 17–24; translation in *Automat. Remote Control*, 49 (1988), pp. 558–565.

- [5] M. A. BERGER AND Y. WANG, *Bounded semigroups of matrices*, Linear Algebra Appl., 166 (1992), pp. 21–27.
- [6] V. D. BLONDEL, J. THEYS, AND A. A. VLADIMIROV, *An elementary counterexample to the finiteness conjecture*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 963–970.
- [7] V. D. BLONDEL, Y. NESTEROV, AND J. THEYS, *On the accuracy of the ellipsoid norm approximation of the joint spectral radius*, Linear Algebra Appl., 394 (2005), pp. 91–107.
- [8] V. D. BLONDEL AND J. N. TSITSIKLIS, *The boundedness of all products of a pair of matrices is undecidable*, Systems Control Lett., 41 (2000), pp. 135–140.
- [9] V. D. BLONDEL AND J. N. TSITSIKLIS, *A survey of computational complexity results in systems and control*, Automat. J. IFAC, 36 (2000), pp. 1249–1274.
- [10] V. D. BLONDEL, S. GAUBERT, AND J. N. TSITSIKLIS, *Approximating the spectral radius of sets of matrices in the max-algebra is NP-hard*, IEEE Trans. Automat. Control, 45 (2000), pp. 1762–1765.
- [11] T. BOUSCH AND J. MAIRESSE, *Asymptotic height optimization for topical IFS, Tetris heaps and the finiteness conjecture*, J. Amer. Math. Soc., 15 (2002), pp. 77–111.
- [12] S. BOYD, L. EL GHAOU, E. FERON, AND V. BALAKRISHNAN, *Linear Matrix Inequalities in System and Control Theory*, SIAM Stud. Appl. Math. 15, SIAM, Philadelphia, 1994.
- [13] R. BRAYTON AND C. TONG, *Constructive stability and asymptotic stability of dynamical systems*, IEEE Trans. Circuits and Systems, 27 (1980), pp. 1121–1130.
- [14] F. COLONIUS AND W. KLIEMANN, *The Dynamics of Control*, Birkhäuser Boston, Boston, MA, 2000.
- [15] I. DAUBECHIES AND J. C. LAGARIAS, *Sets of matrices all infinite products of which converge*, Linear Algebra Appl., 161 (1992), pp. 227–263.
- [16] I. DAUBECHIES AND J. C. LAGARIAS, *Corrigendum/addendum to: Sets of matrices all infinite products of which converge*, Linear Algebra Appl., 327 (2001), pp. 69–83.
- [17] G. GRIPENBERG, *Computing the joint spectral radius*, Linear Algebra Appl., 234 (1996), pp. 43–60.
- [18] N. GUGLIELMI AND M. ZENNARO, *On the zero-stability of variable stepsize multistep methods: The spectral radius approach*, Numer. Math., 88 (2001), pp. 445–458.
- [19] L. GURVITS, *Stability of discrete linear inclusions*, Linear Algebra Appl., 231 (1995), pp. 47–85.
- [20] L. GURVITS, *Stability and observability of discrete linear inclusion – Finite automata approach. Book of abstracts*, in Proceedings of the IEEE International Symposium on the Mathematical Theory of Networks and Systems, Kobe, Japan, 1991, pp. 166–167.
- [21] L. GURVITS, *Stabilities and controllabilities of switched systems (with applications to the quantum systems)*, in Proceedings of the Fifteenth International Symposium on Mathematical Theory of Networks and Systems, University of Notre Dame, Notre Dame, IN, 2002.
- [22] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
- [23] V. S. KOZYAKIN, *Algebraic unsolvability of problem of absolute stability of desynchronized systems*, Automat. Remote Control, 51 (1990), pp. 754–759.
- [24] J. C. LAGARIAS AND Y. WANG, *The finiteness conjecture for the generalized spectral radius of a set of matrices*, Linear Algebra Appl., 214 (1995), pp. 17–42.
- [25] D. LIBERZON, J. P. HESPANHA, AND A. S. MORSE, *Stability of switched systems: A Lie-algebraic condition*, Systems Control Lett., 37 (1999), pp. 117–122.
- [26] M. MAESUMI, *An efficient lower bound for the generalized spectral radius*, Linear Algebra Appl., 240 (1996), pp. 1–7.
- [27] B. E. MOISION, A. ORLITSKY, AND P. H. SIEGEL, *On codes that avoid specified differences*, IEEE Trans. Inform. Theory, 47 (2001), pp. 433–442.
- [28] V. YU. PROTASOV, *The generalized joint spectral radius. A geometric approach*, Izv. Math., 61 (1997), pp. 995–1030.
- [29] G.-C. ROTA AND W. G. STRANG, *A note on the joint spectral radius*, Indag. Math., 22 (1960), pp. 379–381.
- [30] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [31] G. STRANG, *The Joint Spectral Radius*, commentary by Gilbert Strang on paper number 5 in “Collected Works of Gian-Carlo Rota,” 2001.
- [32] J. N. TSITSIKLIS AND V. D. BLONDEL, *The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate*, Math. Control Signals Systems, 10 (1997), pp. 31–40.
- [33] J. N. TSITSIKLIS, *On the stability of asynchronous iterative processes*, Math. Systems Theory, 20 (1987), pp. 137–153.
- [34] J. N. TSITSIKLIS, *The Stability of the Products of a Finite Set of Matrices*, in Open Problems in Communication and Computation, T. M. Cover and B. Gopinath, eds., Springer-Verlag,

- New York, 1987, pp. 161–163.
- [35] A. A. VLADIMIROV, L. ELSNER, AND W.-J. BEYN, *Stability and paracontractivity of discrete linear inclusions*, *Linear Algebra Appl.*, 312 (2000), pp. 125–134.
- [36] F. WIRTH, *The generalized spectral radius and extremal norms*, *Linear Algebra Appl.*, 342 (2002), pp. 17–40.
- [37] F. WIRTH, *On the calculation of real time-varying stability radii*, *Internat. J. Robust Nonlinear Control*, 8 (1998), pp. 1043–1058.